Renewable and Sustainable Energy Reviews 21 (2013) 215–228

Contents lists available at SciVerse ScienceDirect

Renewable and Sustainable Energy Reviews

journal homepage: www.elsevier.com/locate/rser

Wind energy potential assessment for the site of Inner Mongolia in China

Jie Wu [a], Jianzhou Wang [a,*], Dezhong Chi [b]

[a] School of Mathematics & Statistics, Lanzhou University, Lanzhou 730000, China
[b] International Center for Climate and Environment Sciences, Institute of Atmospheric Physics, Chinese Academy of Science, Beijing 100029, China

A R T I C L E I N F O

Article history:
Received 18 September 2012
Received in revised form
27 December 2012
Accepted 29 December 2012
Available online 29 January 2013

Keywords:
Wind speed
Wind energy
Probability density function
Logistic function
Availability factor
Capacity factor

A B S T R A C T

An accurate quantification and characterization of the available wind resources is necessary to optimally design a wind farm. To effectively evaluate the wind energy, studying the wind's statistical characteristics is required. The probability distribution of wind speed is a very important piece of information needed in the assessment of wind energy potential since wind power is proportional to the cube of wind speed. Therefore, choosing a probability function having high goodness of fit with the observation data plays a quite significant role in wind energy assessment. In this study, three probability density functions, i.e., two-parameter Weibull, Logistic and Lognormal are employed to wind speed distribution modeling using data measured at a typical site in Inner Mongolia, China, over the latest three year period from 2009 to 2011. The performance of these three functions is compared so as to select the best one. As one of the most favorable distributions, Weibull function is a most applicable approach in describing wind speed's distribution in many cases. However, though performances of three particle swarm optimization algorithms and 18 differential evolution approaches of estimating the shape parameter estimation in the Weibull function are compared, then the one which performs best is selected to determine the optimal shape parameter to obtain the most accurate shape parameter estimation results. The performance of the Weibull function is worse than the Logistic under the measured wind speed data and the chosen error evaluation criteria. Besides, as compared to the Lognormal function, the Logistic function provides a more adequate result in wind speed distribution modeling. Therefore, in this work, the Logistic function is applied to the consequent wind energy assessment through the availability factor, capacity factor, and turbine efficiency of a wind turbine. Assessment results have shown that it is suitable to build a wind farm in this area.

Contents

* Corresponding author. Tel.: +86 931 8914050; fax: +86 931 8912481.
E-mail addresses: wjz@lzu.edu.cn, 13519619146@126.com (J. Wang).

1364-0321/$ - see front matter
http://dx.doi.org/10.1016/j.rser.2012.12.060

1. Introduction

With the fast development of China's economy over the last few decades, China is becoming a big power in energy consumption. The consumption of coal, petroleum and nature gas constituted 74%, 18.8%, and 4.1% of the total annual primary energy consumption in 2009 [1], respectively, and the corresponding proportions have changed to 68%, 19% and 4.4% in 2010 [2]. Though the total proportion of these three fossil fuels decreased from 96.9% in 2009 to 91.4% in 2010, the total annual primary energy consumption has increased from 2920.28 million tons of standard coal as calorific value calculation in 2009 to 3249.39 million tons in 2010, which means the consumption of these three fossil fuels in 2010 is still larger than that in 2009. As known, fossil fuels are non-renewable resources. Meanwhile, a large amount of greenhouse gases is exhausted in the burning of the fossil fuels, which causes an increasing crisis in global climate warming. Thus, in order to deal with the energy crisis and global climate warming, clean and renewable energy is urgently needed. The "Medium and long term plan of the renewable energy of China" proposed that it is our goal to increase the renewable energy consumption of China up to 15% in 2020 [3]. As one of the cleanest and renewable energy resources, wind energy has been drawing the attention of people, and a large number of wind farms have been built for wind power utilization. Meanwhile, new areas to design wind farms should still be explored. To optimally design a wind farm, an accurate wind energy resource assessment is necessary. It is not only an essential part of the development of wind power utilization, but also provides the investors with the necessary confidence in financial feasibility and mitigating risks [4].

To effectively evaluate the wind power available for a particular site, studying the wind's statistical characteristics is needed. For this reason, a large number of studies have been published concerning the use of a variety of probability density functions (pdfs) to describe wind speed frequency distributions. For example, five distributions including the Weibull, Rayleigh, gamma, lognormal, and inverse Gaussian pdfs were used by Zhou et al. [5] to investigate the wind speed distribution for wind speed data measured from five representative sites in North Dakota. Another two new pdfs named Pearson type V and Burr were introduced by Valerio et al. [6] to describe wind speed frequency distributions in the urban area of Palermo, in the south of Italy. Different from them, the performance of another pdf: the binormal distribution function was tested by Peter and Imre [7] through using ERA-40 surface wind speed data over Europe. Considering that the single Weibull function may not best reflect the distribution of the wind speed, a bimodal Weibull & Weibull pdf was developed by Jaramillo and Borja [8] to analyze the wind speed frequency distribution in La Ventosa, Mexico. In addition, two researchers [9] developed a theoretical approach to the analytical determination of wind speed distributions of four stations in Elazig through application of the maximum entropy principle in 2007, and in 2009, the same two authors [10] represented another theoretical approach of wind speed frequency distributions through applications of a Singly Truncated from below Normal Weibull mixture distribution and a two component mixture Weibull distribution. Furthermore, the minimum cross entropy (MinxEnt) probability density function (pdf) derived from the MinxEnt principle, which covers the maximum entropy principle, was used to determine the diurnal, monthly, seasonal and annual wind speed distributions by Kantar and Usta [11].

Besides, for a specific pdf, methods adopted to determine the parameters in this pdf are various. As one of the most popularly used models due to its ability to fit most accurately the variety of wind speed data measured at different geographical locations in the world [9], the Weibull function is frequently used to the wind speed distribution modeling. Methods applied to Weibull parameters estimation commonly including: the graphical method [12], the maximum likelihood method [13], the moment estimation and quartiles [14], as well as the modified maximum likelihood and energy pattern factor method [15], etc.

In this paper, three probability density functions including the Weibull, Logistic and Lognormal functions are employed to wind speed distribution modeling. To obtain the most accurate shape parameter value in Weibull function, three particle swarm optimization algorithms and 18 differential evolution approaches are adopted, and the one with the best performance is selected to determine the final optimal shape parameter. Then the performance of these three probability density functions is evaluated according to four error evaluation criteria by giving them different weight, the optimal function in modeling the wind speed distribution is chosen to assess the wind energy of this area.

The rest of this paper is organized as follows: Section 2 presents the wind energy potential assessment review. In Section 3, the extrapolated data in the hub height are presented as well as a brief analysis to the data. Theory of the two-parameter Weibull function is introduced in Section 4, and the objective function which will be used in the subsequent parameter optimization process is also exhibited. Then the algorithms of three particle swarm optimization and 18 differential evolution approaches are presented in Section 5. Section 6 gives the final optimal results in estimating the shape parameter in the Weibull function. The performance of two other probability density functions are compared with the Weibull function under four error evaluation criteria in Section 7, and the best function

selected in Section 6 is applied to the wind energy assessment in Section 8. Finally, Section 9 concludes the findings of this paper.

2. Wind energy potential assessment review

2.1. Assessment related to the Weibull distribution

Due to the superiority performance of the Weibull distribution in describing the wind speed distribution, many wind energy potential assessment works have been carried out with the help of the Weibull distribution.

Among these works, periods of the wind speed data covering are various. Weisser [16] utilized the Weibull density function to estimate the wind energy potential in Grenada, West Indies by the wind speed data from 1996 to 1997. A data source consisting of five years (1998–2002) of hourly mean wind data was adopted to determine the wind energy potential in Maden-Elazig, Turkey [17] with the assistance of the Weibull distribution. The same distribution was also used by Keyhani et al. [18] to find out the wind energy potential of the capital of Iran, Tehran according to the statistical data of eleven years' wind speed measurements. Ohunakin and Akinnawonu [19] examined the wind energy potential in Jos, Nigeria using the wind speed data during 1971–2007 subjected to the Weibull analysis as well. Moreover, Chang et al. [20] analyzed the wind characteristics in Taiwan based on an even long period from 1961 to 1999 so as to present an assessment of the wind energy potential in this area.

In addition, the characteristics of the wind speed data at different heights can also be described by the Weibull distribution so as to provide a reliable wind energy potential assessment. The possible wind energy potential of Kütahya, Turkey at the single 30 m height was investigated by Köse [21] according to the Weibull distribution. Ullah et al. [22] employed the Weibull distribution to provide an evaluation of wind energy potential in Kati Bandar, Pakistan at three different heights, i.e., 10 m, 30 m and 50 m. The Weibull distribution was also employed by Mirhosseini et al. [23] to assess the wind energy potential locations in province of Semnan in Iran with the wind speed data at 10 m, 30 m and 40 m height. And based on the wind data collected at the heights of 10, 40, 50 and 60 m, Đurišić and Mikulović [24] carried out an assessment on wind energy in the South Banat region, Serbia.

As a special form of the Weibull function, the Rayleigh distribution function has also been frequently used to assess the wind energy potential. A Rayleigh distribution was adopted for defining the distribution of wind velocity in terms of its probability density and cumulative distribution functions by Mathew et al. [25] so as to present a wind energy potential assessment of eight sites in Kerala, India. Rayleigh parameter at a height of 30 m has been estimated and used by Amar et al. [26] to assess the wind energy potential of Sidi Daoud, Tunisia.

2.2. Assessment related to other approaches

Apart from the probability density functions, some other approaches have been applied to the wind energy potential assessment as well. Onat and Ersoz [27] analyzed the wind energy potential of three regions in Turkey with the help of the Wind Atlas Analysis and Application Program software, the same software was also used by Durak and Şen [28] to assess the wind power potential in Turkey and Akhisar. Hossain et al. [29] assessed the wind energy potential for India by an innovative approach using the geographical information system (GIS) platform. A wind resource inventory created by the GIS data supplied with the wind maps which was developed with a meso-scale wind energy model coupled with wind data was presented by Acker et al. [30] to assess the wind energy potential in the state of Arizona as well. And Thiaw et al. [31] used the multi-layer perceptron type artificial neural networks to determine the wind energy potential on the Dakar site. Besides, the fifth-generation meso-scale Model was used in conjunction with an h-adaptive finite element model by Pepper and Wang [32] to assess the wind energy potential of Nevada, and an assessment of long-term wind energy potential at the demonstration offshore wind farm in Korea was estimated by Oh et al. [33] by using the Measure–Correlate–Predict technique.

3. Site representation and brief analysis

3.1. Site representation

Wind speed data used in this paper is measured from 2009 to 2011 at one of the sites in Inner Mongolia, China, with the latitude of 42.5°N and longitude of 115°E. The data are sampled at the standard height 10 m with a temporal resolution of 6 h, so there are four data for one day. The cut-in, rated and cut-off speeds of the wind turbine S 43/600 in this area are 3, 15 and 24 m/s, respectively, and the rated power is 600 kW. Rotor's diameter is 43 m, which has a blade sweep area of 1452.2 m^2, and the hub height is 50 m above the ground level.

Wind speed values sampled at the standard height are first extrapolated to the ones at the hub height using the following empirical power law:

$$\frac{v}{v_1} = \left(\frac{h}{10}\right)^{\alpha}, \tag{1}$$

where v_1 and v are the wind speeds at the standard height 10 m and at the hub height h m, respectively, and α is the roughness factor, in this analysis, a typical value of 0.14 is adopted. By substituting $h=50$ and the measured wind speed data at the standard height 10 m to Eq. (1), wind speed values at the hub height are obtained. Fig. 1 shows the variation trend of the wind speed throughout the year.

As demonstrated in Fig. 1, the highest and lowest wind speeds in this area are about 18 and 2 m/s, respectively, which appear in February and May. In addition, except for May, June is also one of the low wind speed period.

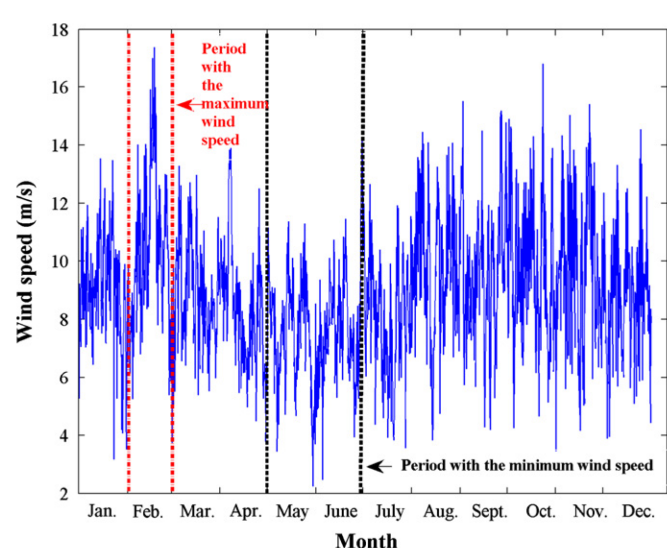


Fig. 1. The variation trend of the wind speed throughout the year.

Table 1
Brief analysis of the wind speed characteristics.

Characteristic	Period												
	January	February	March	April	May	June	July	August	September	October	November	December	Yearly
Mean speed (m/s)	8.9439	10.2558	8.6957	8.2298	7.4566	7.5508	7.8993	9.6786	9.5821	9.5605	9.6972	8.7084	8.8444
Standard deviation (m/s)	1.9704	2.6717	1.8471	2.0910	1.8862	1.9052	2.0251	2.3573	2.3878	2.5014	2.6589	2.0899	2.3738
Skewness	−0.512	0.374	0.198	0.670	−0.288	0.712	0.027	−0.071	0.178	0.373	−0.049	0.196	0.338
Kurtosis	0.526	−0.248	−0.188	0.530	−0.252	1.144	−0.407	−0.379	0.188	−0.361	−0.674	0.003	0.067

3.2. Brief analysis to the wind speed data

In this section, the wind speed data at the hub height are analyzed from the following three aspects: the central tendency, the discrete degree and the distributional pattern.

3.2.1. Introduction to the statistics

3.2.1.1. Statistics used to describe the central tendency. The central tendency refers to the tendency of a set of data which draws close to a central value. Statistics used to describe the central tendency reflects the representative value or the center value of a data sequence. The mean value of a data sequence is usually employed to describe the central tendency, which is defined as follows:

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i, \tag{2}$$

where $\bar{x}$ is the mean value, $\{x_i\}_{i=1}^{n}$ is the data sequence and n is the number of the data contained in this data sequence.

3.2.1.2. Statistics used to describe the discrete degree. Using only the mean of a data sequence cannot depict the property of the data completely. Apart from the central value, the 'distance' of all data relative to the central value should also be investigated. If the data closely gather around the central value, the discrete degree of the data is small, which means the central value is the representative of the data sequence; Otherwise, if the data are loosely distributed around the central value, the discrete degree of the data is large and the central value cannot represent all of the data well in this case. Therefore, the central value and the discrete degree of the data should be given together to give a comprehensive description to the data sequence. The discrete degree of the data sequence relative to the mean value can be described by the Standard deviation (*Std*) value expressed as follows:

$$Std = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n} (x_i - \bar{x})^2}, \tag{3}$$

where $\bar{x}$ is defined by Eq. (2).

3.2.1.3. Statistics used to describe the distributional pattern. To grasp the characteristics of the data better, the distributional pattern of the data should also be surveyed. The distributional pattern can be researched by two statistics: *Skewness* and *Kurtosis*, the expression of these two statistics are shown in the following equations:

$$Skewness = \frac{1}{n-1}\sum_{i=1}^{n} (x_i - \bar{x})^3 / Std^3, \tag{4}$$

$$Kurtosis = \frac{1}{n-1}\sum_{i=1}^{n} (x_i - \bar{x})^4 / Std^4 - 3, \tag{5}$$

where $\bar{x}$ and *Std* are defined by Eqs. (2) and (3), respectively. *Skewness* is a description of the symmetrical characteristic of the data, where $Skewness = 0$ means the distributions of the data is symmetric, or else, it is not symmetric, and the larger the absolute of the *Skewness* value, the more skewness of the data. *Kurtosis* is a statistic employed to describe the steep degree of the data, where $Kurtosis = 0$ representing the steep degree of the data is the same as the standard normal distribution, $Kurtosis > 0$ means the distribution of the data is steeper than the standard normal distribution while $Kurtosis < 0$ reflects the distribution of the data is less steep than the standard normal distribution.

3.2.2. Analysis results to the wind speed data

Applying the above statistics introduced in Section 3.2.1 to the wind speed data, results listed in Table 1 are obtained.

As demonstrated in Table 1, the mean speed in February is the maximum, while the one in May is the minimum. Wind speed data in November present the highest degree far away from the mean value for the largest standard deviation value it reached, and the discrete degree of the wind speed in March is the lowest. Besides, the skewness of the wind speed data in June is the most severe, and wind speeds in January, April, June, September, December and the whole year are steeper than the standard normal distribution (for the *Kurtosis* values of these months are larger than 0), while the wind speed data in other months are less steep than the standard normal distribution.

Surveying the wind speed characteristics and distribution according to these statistics is not enough. More detailed information should be acquired by analyzing the wind speed distribution and fitting it through some distribution functions, such as the two-parameter Weibull model, the Logistic and the Lognormal distributions, etc.

4. Two-parameter Weibull model

4.1. Weibull function

Weibull probability density function (pdf) is expressed as

$$f(v) = \left(\frac{k}{c}\right)\left(\frac{v}{c}\right)^{k-1} \exp\left[-\left(\frac{v}{c}\right)^k\right], \tag{6}$$

where v, k, and c are the wind speed, the dimensionless shape parameter, and the scale parameter with the same unit as the wind speed, respectively. Based on the Weibull pdf, Weibull cumulative distribution function (cdf) can be calculated by the integration:

$$\begin{aligned} F(v) &= \int_0^v f(s)ds = \int_0^v \left(\frac{k}{c}\right)\left(\frac{s}{c}\right)^{k-1} \exp\left[-\left(\frac{s}{c}\right)^k\right] ds \\ &= -\int_0^v d\exp\left[-\left(\frac{s}{c}\right)^k\right] = -\exp\left[-\left(\frac{s}{c}\right)^k\right]\Bigg|_{s=0}^{s=v} \\ &= 1 - \exp\left[-\left(\frac{v}{c}\right)^k\right]. \end{aligned} \tag{7}$$

Table 2
Simulation results by the three PSO algorithms.

Algorithm	Item	**Minimal iteration when the termination criterion is the fitness level is 1×10^{-5}**												
		January	February	March	April	May	June	July	August	September.	October	November	December	Yearly
BPSO	Iteration	29	28	50	43	31	54	88	38	50	42	43	42	50
	k	5.2172	4.3412	5.4290	4.4622	4.4836	4.4965	4.4184	4.6744	4.5584	4.3205	4.1024	4.7513	4.2007
	Fitness ($\times10^{-6}$)	6.0542	6.3340	7.1916	1.7391	3.6284	1.0514	1.8565	3.0458	1.8726	5.0785	4.5689	9.2880	8.1431
WPSO	Iteration	21	24	26	17	19	18	22	17	20	16	14	24	15
	k	5.2175	4.3408	5.4290	4.4622	4.4835	4.4961	4.4180	4.6749	4.5583	4.32	4.1027	4.7505	4.2006
	Fitness ($\times10^{-6}$)	1.7934	3.7623	6.5053	3.3368	5.4070	9.6330	7.2239	9.3189	0.1790	7.7823	3.7711	8.7104	4.5387
QPSO	–[a]	–[a]	–[a]	–[a]	–[a]	–[a]	–[a]	–[a]	–[a]	–[a]	–[a]	–[a]	–[a]	–[a]
		Fitness value when the termination criterion is the maximum iteration number equals to 50												
BPSO	k	5.2174	4.3408	5.4296	4.4620	4.4838	4.4964	4.4183	4.6742	4.5585	4.3202	4.1026	4.7505	4.2005
	Fitness ($\times10^{-6}$)	2.8315	2.9134	1.8608	2.5735	0.5046	1.9653	0.2005	7.3031	4.4822	3.3872	0.3628	7.6569	2.0268
WPSO	k	5.2176	4.3409	5.4295	4.4621	4.4838	4.4964	4.4183	4.6745	4.5583	4.3203	4.1025	4.7509	4.2004
	Fitness ($\times10^{-8}$)	0.1586	0.7320	1.5120	1.0433	0.1898	0.0651	0.0439	0.0437	0.4090	0.0703	0.2473	0.0121	0.0403
QPSO	k	5.2176	4.3409	5.4295	4.4621	4.4838	4.4964	4.4183	4.6745	4.5583	4.3203	4.1025	4.7509	4.2004
	Fitness ($\times10^{-10}$)	0.3651	0.0872	0.0926	0.0401	0.2381	0.0183	0.1014	0.8759	0.0230	0.0348	0.0541	0.0002	1.7107

[a] Incalculable value.

4.2. Objective function

Based on the Weibull pdf, the mean $\overline{v}$ and standard deviation (σ) can be obtained:

$$\begin{aligned}\overline{v} &= \int_0^{+\infty} v f(v)\mathrm{d}v \\ &= \int_0^{+\infty} v\left(\frac{k}{c}\right)\left(\frac{v}{c}\right)^{k-1}\exp\left[-\left(\frac{v}{c}\right)^k\right]\mathrm{d}v \\ &= \int_0^{+\infty} v\exp\left[-\left(\frac{v}{c}\right)^k\right]\mathrm{d}\left(\frac{v}{c}\right)^k \\ &= \int_0^{+\infty} c\left[\left(\frac{v}{c}\right)^k\right]^{(1+1/k)-1}\exp\left[-\left(\frac{v}{c}\right)^k\right]\mathrm{d}\left(\frac{v}{c}\right)^k \\ &= c\int_0^{+\infty}\left[\left(\frac{v}{c}\right)^k\right]^{(1+1/k)-1}\exp\left[-\left(\frac{v}{c}\right)^k\right]\mathrm{d}\left(\frac{v}{c}\right)^k \\ &= c\Gamma\left(1+\frac{1}{k}\right), \end{aligned} \tag{8}$$

$$\begin{aligned}\sigma^2 &= E(v-\overline{v})^2 = Ev^2-(Ev)^2 \\ &= \int_0^{+\infty} v^2\left(\frac{k}{c}\right)\left(\frac{v}{c}\right)^{k-1}\exp\left[-\left(\frac{v}{c}\right)^k\right]\mathrm{d}v - c^2\Gamma^2\left(1+\frac{1}{k}\right) \\ &= \int_0^{+\infty} c^2\left[\left(\frac{v}{c}\right)^k\right]^{(1+2/k)-1}\exp\left[-\left(\frac{v}{c}\right)^k\right]\mathrm{d}\left(\frac{v}{c}\right)^k - c^2\Gamma^2\left(1+\frac{1}{k}\right) \\ &= c^2\Gamma\left(1+\frac{2}{k}\right) - c^2\Gamma^2\left(1+\frac{1}{k}\right). \end{aligned} \tag{9}$$

To search the Weibull parameters by using the intelligent parameter optimization approaches, the residual value ε defined as below is used as the objective function in this paper just as Liu et al. did in [34]:

$$\varepsilon = \frac{\sigma^2}{\overline{v}^2} - \frac{\Gamma(1+2/k)-\Gamma^2(1+1/k)}{\Gamma^2(1+1/k)}. \tag{10}$$

If the objective function converges to a predefined tolerance value then an acceptable shape parameter k is obtained; furthermore, the scale parameter c can be determined using Eq. (8). In the process of the parameter searching, when a given maximum iteration number has been produced or a satisfactory fitness level is reached, the final value of the parameter is determined and the algorithm is terminated.

5. Parameter optimization approaches

To search the optimal shape parameter in the Weibull function, the following intelligent parameter optimization approaches are employed.

5.1. Three particle swarm optimization (PSO) algorithms

5.1.1. Basic PSO (BPSO)

PSO is a population based optimization tool, where the system is initialized with a population of random particles and the algorithm searches for optima by updating generations. Suppose that the search space is n-dimensional, the ith particle is represented by an n-dimensional vector $X_i=(x_{i1},x_{i2},\ldots,x_{in})$ and the velocity of this particle is represented as $V_i=(v_{i1},v_{i2},\ldots,v_{in})$. According to the objective function, the fitness of each particle can be evaluated. Record the fitness of the personal best position $P_i=(p_{i1},p_{i2},\ldots,p_{in})$, which is the best previously visited position of the particle i, as f_{pbest}. Meanwhile, the position of the best individual of the whole swarm is marked as the global best position $P_g=(P_{g1},P_{g2},\ldots,P_{gn})$, and the objective function corresponding to P_g is f_{gbest}. At each step, the velocity and position of the ith particle are updated according to the following two equations until a user-defined stopping limit is reached:

$$V_i = \omega * V_i + c_1 * r_1 * (P_i - X_i) + c_2 * r_2 * (G - X_i), \tag{11}$$

$$X_i = X_i + V_i, \tag{12}$$

where ω is called the inertia weight, by which the impact of previous velocity of particle on its current one can be controlled, r_1 and r_2 are independently uniformly distributed random variables with range (0, 1) and c_1 and c_2 are positive constant parameters called acceleration coefficients which control the maximum step size. In the BPSO algorithm of this paper, ω is a random value in the range (0, 1).

5.1.2. PSO with self-adaptive stochastic inertia weight (WPSO)

As exhibited in Eq. (11), the velocity of particle is influenced by the inertia weight ω. Different from the BPSO, in the WPSO algorithm, the inertia weight ω is defined as values in different ranges according to the change of f_{gbest}:

(a) If f_{gbest} is unchanged, ω is defined as a random value in the range (0.5, 1) to increase the search intensity.

(b) Otherwise, if f_{gbest} is changed, ω is defined as the same as in the BPSO algorithm, i.e., it is defined as a random value in the range (0, 1) [35].

5.1.3. Quantum-behaved particle swarm optimization (QPSO)

QPSO algorithm is a new global optimization algorithm based on the PSO algorithm. Different from the PSO algorithm, only the location information for the particles is given in QPSO algorithm, the location is updated by the following three equations:

$$p_{id} = \phi * pbest_{id} + (1-\phi) * gbest_d, \tag{13}$$

$$mbest = \frac{1}{M}\sum_{i=1}^{M} pbest_i, \tag{14}$$

$$x_{id} = p_{id} \pm \alpha * |mbest_d - x_{id}| * \ln(1/u), \tag{15}$$

where ϕ and u are random values in the range (0,1), thus p_{id} is a random position between *pbest* and *gbest*, *M* is the size of the particles, and α is called the expansion-contraction factor. As shown in Eq. (14), *mbest* is the center of the current optimal location in all individual particles. Parameter α in Eq. (15) is used to control the convergence speed of the algorithm. It is the sole parameter needed to determine and control, usually it is determined as the linear function of the iterations, that is

$$\alpha = (\alpha_{\max} - \alpha_{\min}) * (iteration - iter)/iteration + \alpha_{\min} \tag{16}$$

where *iter* is the current iterations, *iteration* is the total iterations, $\alpha_{\max}$ and $\alpha_{\min}$ are two positive constants, which are usually set as 1.0 and 0.5 [36].

5.2. Differential evolution

In the algorithm of the differential evolution (DE), a set of *D* optimization parameters is called an individual, which is represented by a *D*-dimensional parameter vector, and the population consists of *NP* parameter vectors X_i^G (i=1,2,…,*NP* for each generation *G*). The basic process of DE includes three steps: mutation, crossover and selection, which are described as follows:

Step 1, *Mutation*: For each target vector X_i^G (i=1,2,…,*NP*), the mutation operation is conducted according to Eq. (17) so as to generate a mutant vector V_i^{G+1}:

$$v_i^{G+1} = x_{r1}^G + F * (x_{r2}^G - x_{r3}^G),\ r1 \neq r2 \neq r3 \neq i, \tag{17}$$

where $r1$, $r2$, $r3$ are integer indexes randomly chosen from {1,2,…,*NP*}, *F* is called the mutation factor and is taken from the interval [0,1]. Actually, it is a factor which controls the amplification of the difference between two individuals. In this paper, this mutation method is called Strategy 1. In addition, eight other mutation approaches which are represented as follows are employed to optimal parameter searching in this paper:

Strategy 2: $v_i^{G+1} = x_i^G + F_1 * (x_{best}^G - x_i^G) + F_2 * (x_{r2}^G - x_{r3}^G)$;
Strategy 3: $v_i^{G+1} = x_{best}^G + (x_{r1}^G - x_{r2}^G) * ((1-0.9999) * rand + F)$;
Strategy 4: $v_i^{G+1} = x_{r1}^G + F_1 * (x_{r2}^G - x_{r3}^G)$, $F_1 = (1-F)^* rand + F$, and F_1 are equal for all of the parameters needed to estimate;
Strategy 5: $v_i^{G+1} = x_{r1}^G + F_1 * (x_{r2}^G - x_{r3}^G)$, $F_1 = (1-F)^* rand + F$;
Strategy 6: $v_i^{G+1} = x_i^G + F * (x_{r2}^G - x_{r3}^G)$;
Strategy 7: $v_i^{G+1} = x_{r1}^G + F * (x_{r2}^G - x_{r3}^G + x_{r4}^G - x_{r5}^G)$;
Strategy 8: $v_i^{G+1} = x_{best}^G + F * (x_{r2}^G - x_{r3}^G + x_{r4}^G - x_{r5}^G)$;
Strategy 9:

$$v_i^{G+1} = \begin{cases} x_{r1}^G + F * (x_{r2}^G - x_{r3}^G) & \text{if } rand < 0.5, \\ x_{r1}^G + 0.5 * (F+1) * (x_{r1}^G + x_{r2}^G - 2 * x_{r3}^G) & \text{if } rand \geq 0.5, \end{cases}$$

where $r4$ and $r5$ are integer indexes randomly chosen from 1,2,…,*NP* as well, x_i^G and x_{best}^G are the *i*th and the best individuals in generation *G*, respectively.

Step 2, *Crossover:* In order to increase the diversity of the perturbed parameter vectors, crossover is introduced. The target vector is mixed with the mutated vector, using the following scheme, to yield the trial vector $U_i^{G+1} = (u_{1i}^{G+1}, u_{2i}^{G+1}, \ldots, u_{Di}^{G+1})$, that is

$$u_{ji}^{G+1} = \begin{cases} v_{ji}^{G+1} & \text{if } rand(j) \leq CR \text{ or } j = rnb(i) \\ x_{ji}^G & \text{otherwise} \end{cases} \quad j = 1,2,\ldots,D \tag{18}$$

where *rand*(*j*) is the *j*th evaluation of a uniform random number generator between [0,1]. *CR* is the crossover constant between [0,1] which has to be determined by the user. In Eq. (18), *rnb*(*i*) is a randomly chosen index from 1,2,…,*D* which ensures that U_i^{G+1} gets at least one parameter from V_i^{G+1}. Otherwise, no new parent vector would be produced and the population would not alter. This is called the binomial crossover. Besides this, another crossover approach named the exponential crossover is also a commonly used crossover method, in which the trial vector U_i^{G+1} is described as

$$u_{ji}^{G+1} = \begin{cases} v_{ji}^{G+1} & \text{if } j \in \{k, \langle k+1 \rangle_n, \ldots, \langle k+L-1 \rangle_n\}, \\ x_{ji}^G & \text{otherwise}, \end{cases} \quad j = 1,2,\ldots,D, \tag{19}$$

where $k \in \{1,2,\ldots,n\}$ is a random index, *L* is a random value in $\{1,2,\ldots,n\}$ and $\langle j \rangle_n$ is *j* if $j \leq n$ and $j-n$ if $j > n$.

Step 3, *Selection*: Assuming that it is the goal of the objective function to obtain the minimal value of one variable, the selection operation is carried out according to the following rule:

$$X_i^{G+1} = \begin{cases} U_i^{G+1} & \text{if } f\left(U_i^{G+1}\right) \leq f\left(X_i^G\right), \\ X_i^G & \text{otherwise}, \end{cases} \tag{20}$$

That is, if vector U_i^{G+1} yields a smaller fitness value than X_i^G, then X_i^{G+1} is set to U_i^{G+1}; otherwise, the old value X_i^G is retained. The purpose of this step is to guarantee all the individuals of the next generation are as good as or better than their counterparts in the current generation.

6. Shape parameter searching results

6.1. Simulation results by the above three PSO algorithms

Based on the three PSO algorithms introduced in Section 5.1, the shape parameter (k) is estimated under two termination criterions: one is the fitness level equals 1×10^{-5} and the other is the maximum iteration number equals 50. Simulation results by the three PSO algorithms are listed in Table 2.

Note that in the QPSO algorithm, the total iterations, i.e., the maximum iteration number, must be given before searching the optimal parameters, however, if the termination criterion is the fitness level other than the maximum iteration number, this algorithm cannot be adopted any more. Thus, in Table 2 corresponding values are incalculable for QPSO when the fitness level (1×10^{-5}) is set as the iteration stopping limit. As shown in Table 2, though the optimal shape parameters searched by the three PSO algorithm are nearly the same, the WPSO and QPSO algorithms both obtain better performance in searching the optimal shape parameter (k) than the BPSO: fewer iterations or more accurate fitness value. Besides, when the termination criterion is the maximum iteration number equals 50, the QPSO

obtains the highest fitness accuracy for most of the months, however, the best performance is not suitable for all of the months, thus the best algorithm should be chosen according to the actual requirements. For example, to receive the best parameter optimization performance for December, it is better to choose the WPSO model to reach the minimum fitness value.

6.2. *Shape parameter searching results by the DE approach*

In this section, the nine mutation strategies mentioned in Section 5.2.1 and the two crossover approaches introduced in Section 5.2.2 are combined to generate 18 different DE methods. As shown in Eq. (18), the crossover constant CR is determined by the users, thus in Table 3 the performances of these methods with two different crossover constants are surveyed: $CR=1$ and $CR=0.5$. Generally, as seen from Table 3, a smaller crossover constant needs larger iteration number to reach the fitness level. This phenomenon is best explained according to the minimum iteration number reached by the strategy 7-exp in the case of two different crossover constants, though their comment goal is to reach the fitness level (1×10^{-5}). Thus, a bigger crossover constant can be selected to search the optimal parameter more quickly. However, though the crossover constant 1 is two times to 0.5, the corresponding minimum iterations under these two crossover constants do not have this relationship.

Similar to the three PSO algorithms, the performance of these 18 DE approaches when the termination criterion is set as the maximum iteration number equals 50 for each month and the whole year is researched. The results are shown in Table 4. Here, only the results in the case of $CR=1$ are shown. As seen, Strategy 2 and Strategy 3 obtain more accurate fitness values than other strategies as well.

6.3. *Final optimal shape parameter results*

In this paper, the tolerance value of the objective function is defined to be 1×10^{-5} so as to obtain more accurate results than it did (1×10^{-3}) in Ref. [34]. In addition, the optimal parameter in each month and the whole year is selected by comparing the three PSO and the 18 DE algorithms. The final optimal parameter for each time stage is selected by comparing the minimum iteration number. The approach which obtains the smallest iteration number is determined to be the best method and then the shape parameter is estimated by this approach. The shape parameter in each month and the whole year determined by which algorithm are plotted in Fig. 2, among which, '[*a*]', '[*b*]', '[*c*]', '[*d*]', '[*e*]' and '[*f*]' mean this optimal shape parameter is searched by the Strategy 2-Bin in the case of $CR=0.5$, Strategy 2-Bin in the case of $CR=1$, Strategy 3-Bin in the case of $CR=1$, Strategy 1-exp in the case of $CR=1$, Strategy 9-Bin in the case of $CR=1$ and Strategy 1-Bin in the case of $CR=1$, respectively. Scale parameters determined by the corresponding shape parameters and the minimum iteration numbers are shown in Fig. 2 as well.

As presented in Fig. 2, the shape parameters searched by the Strategy 2-Bin in the case of $CR=1$ and the Strategy 3-Bin in the case of $CR=1$ are the most, and all of the optimal shape parameters are searched by the DE approaches. Besides, all of the shape parameters vary near 5, and the scale parameter line is located above the shape parameter line. The minimum iteration number in searching the optimal shape parameters for the yearly wind speed data is no more than 5, which is also the smallest in all of the shape parameters searching process. Fig. 3 shows the complete procedure for wind energy assessment, from which it can be seen that the optimal shape parameter determination is a necessary part of the best probability density function selection, so as to provide an accurate wind energy assessment.

Table 3
Minimum iteration numbers by 18 different DE algorithms when the termination criterion is the fitness level equals 1×10^{-5}.

Strategy		$CR=1$													$CR=0.5$												
		January	February	March	April	May	June	July	August	September	October	November	December	Yearly	January	February	March	April	May	June	July	August	September	October	November	December	Yearly
Strategy 1	Bin	20	19	14	18	16	18	20	16	19	23	9	16	3	5	23	39	42	41	37	33	32	31	37	30	26	37
	Exp	20	16	18	9	16	14	20	13	8	14	8	17	18	70	97	28	89	86	58	106	29	23	35	54	73	56
Strategy 2	Bin	8	5	9	10	7	12	16	13	15	7	9	16	9	4	15	23	25	25	32	24	22	24	23	40	29	34
	Exp	11	13	12	10	13	14	12	15	11	16	10	14	9	46	6	25	70	60	49	37	31	37	55	43	61	59
Strategy 3	Bin	13	14	9	10	11	15	6	15	11	18	7	14	13	27	33	15	17	25	29	28	29	30	25	22	30	14
	Exp	11	12	10	13	8	15	8	15	11	10	20	15	16	6	19	37	32	30	29	33	35	35	51	34	33	21
Strategy 4	Bin	21	16	16	21	17	18	14	16	23	13	16	17	14	39	39	50	40	41	29	26	37	34	33	26	38	23
	Exp	15	22	17	18	21	18	18	13	15	14	9	21	21	76	89	55	100	61	48	66	114	138	71	96	48	119
Strategy 5	Bin	22	20	20	15	18	23	12	17	20	8	20	19	16	37	38	42	47	46	43	37	38	40	41	29	46	35
	Exp	22	15	21	22	12	14	20	21	20	15	20	23	20	48	77	37	51	40	98	47	34	78	69	63	87	62
Strategy 6	Bin	23	17	14	32	29	29	26	14	28	29	19	23	26	41	42	22	65	59	37	39	59	38	63	72	50	22
	Exp	19	35	27	20	24	27	22	16	29	28	25	32	26	93	62	67	66	63	53	88	75	65	106	53	70	86
Strategy 7	Bin	28	22	27	11	21	19	23	15	22	25	19	31	25	46	64	40	38	38	37	35	39	35	40	44	36	64
	Exp	25	20	27	18	18	25	9	24	20	20	12	24	21	428	213	281	170	92	73	145	233	153	460	96	190	212
Strategy 8	Bin	25	32	63	27	24	21	23	29	24	20	24	26	21	46	57	31	43	61	70	51	63	61	42	51	48	68
	Exp	28	22	17	14	21	23	27	45	24	25	20	32	30	98	103	44	100	76	170	165	431	153	283	79	116	202
Strategy 9	Bin	18	20	15	19	18	21	21	10	19	15	21	17	17	36	33	46	23	39	34	41	39	27	37	28	35	28
	Exp	17	20	15	16	21	23	21	14	21	20	16	17	20	67	32	64	41	68	89	80	15	99	103	64	107	65

Table 4
Simulation results by DE when the termination criterion is the maximum iteration number equals 50.

Strategy		Fitness value												
		January	February	March	April	May	June	July	August	September	October	November	December	Yearly
Strategy 1	Bin ($\times 10^{-9}$)	0.1573	0.0028	0.1070	0.0157	1.2493	0.0374	0.2927	0.1141	2.2147	0.0138	0.1207	0.0503	0.3497
	Exp ($\times 10^{-9}$)	0.0288	0.7259	0.0281	0.0176	0.5038	0.1216	0.3234	0.7072	0.5759	0.0115	0.1450	0.2090	0.0595
Strategy 2	Bin ($\times 10^{-12}$)	0.8157	0.3215	0.0147	0.5550	0.0582	0.3916	0.0456	0.1642	0.1093	0.1437	0.5338	0.0271	0.2491
	Exp ($\times 10^{-12}$)	0.0692	0.3393	0.9403	0.5266	0.1646	7.1883	0.3278	0.0291	0.0159	7.4262	0.4241	1.8338	0.0245
Strategy 3	Bin ($\times 10^{-12}$)	0.0818	0.1382	0.1490	0.0607	0.0130	0.0096	0.0542	0.1126	0.1200	0.7047	0.1018	0.0163	5.9767
	Exp ($\times 10^{-12}$)	1.0958	0.1602	1.6932	0.6109	0.0257	1.5624	0.0738	0.6123	0.1136	0.0173	0.0511	0.0482	0.2150
Strategy 4	Bin ($\times 10^{-9}$)	0.2092	0.0524	0.2356	0.1324	0.4212	0.1525	0.0096	0.2009	3.6355	0.1204	0.7823	0.2187	0.2330
	Exp ($\times 10^{-9}$)	0.5701	0.1092	0.6607	0.0023	0.2526	0.0122	0.2946	0.0526	0.8931	0.0829	0.4739	0.3023	0.0827
Strategy 5	Bin ($\times 10^{-8}$)	0.0116	0.0104	0.0697	0.0023	0.0199	2.7678	0.0758	0.0012	0.2862	0.0450	0.0149	0.0423	0.0008
	Exp ($\times 10^{-8}$)	0.0073	0.0144	0.0141	0.1230	0.0123	0.0422	0.0015	0.0168	0.0002	0.2261	0.1335	0.0013	0.0004
Strategy 6	Bin ($\times 10^{-7}$)	0.1539	0.0781	2.2850	5.0708	0.9624	0.0977	0.1522	2.0323	0.8490	1.9373	0.1271	0.0386	1.2156
	Exp ($\times 10^{-7}$)	1.3920	0.2046	0.0901	0.3051	0.6177	1.8480	0.4300	3.0977	0.5155	0.3120	0.5847	0.0331	0.2196
Strategy 7	Bin ($\times 10^{-7}$)	0.0445	0.0184	0.0898	0.0143	0.0315	0.0147	0.0118	0.0086	0.0169	0.0548	0.0175	0.0048	0.0029
	Exp ($\times 10^{-7}$)	0.1461	0.0059	1.3403	0.0312	0.1595	0.0340	0.0469	0.0007	0.0195	0.0950	0.0085	0.0438	0.0123
Strategy 8	Bin ($\times 10^{-6}$)	0.1829	0.0450	0.0136	1.1316	0.0631	0.0154	0.0563	3.6629	0.9170	0.0681	0.4209	0.6739	0.0070
	Exp ($\times 10^{-6}$)	0.2409	0.5098	0.0708	0.0201	0.0026	0.4208	0.5867	0.0429	0.1287	0.1023	0.0531	0.0260	0.0071
Strategy 9	Bin ($\times 10^{-9}$)	0.7274	0.0992	9.8622	0.0012	0.0265	0.8364	0.1521	0.0202	0.0002	0.8163	0.0671	1.6552	0.0211
	Exp ($\times 10^{-9}$)	0.5127	0.0245	0.0034	0.1061	0.7093	0.0198	0.4501	0.0067	0.0018	0.4457	0.0968	0.267	0.0125

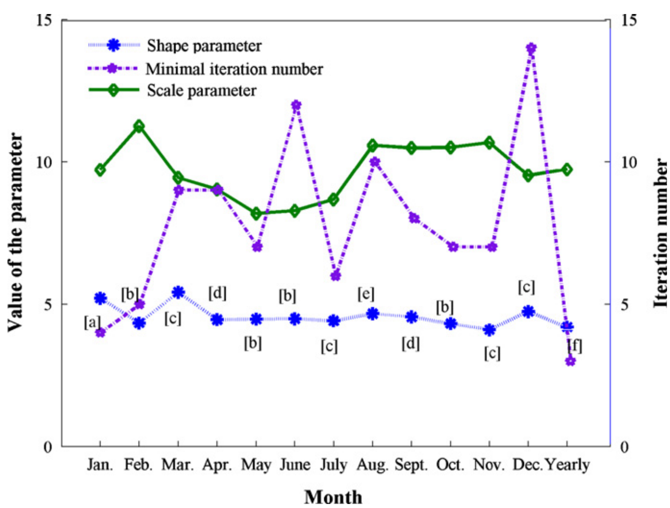


Fig. 2. Shape, scale parameter and minimum iteration number during different time periods.

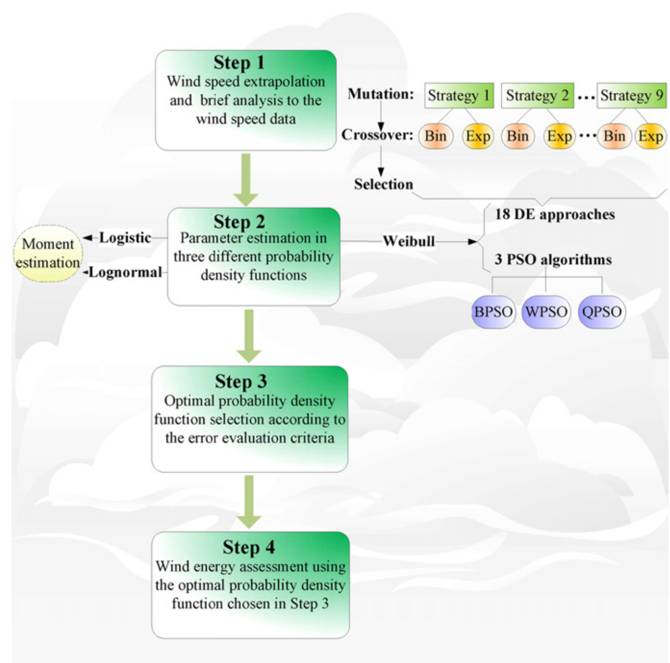


Fig. 3. Flowchart of wind energy assessment.

7. Optimal probability density function selection

7.1. Two other distributions used to compare

To evaluate the performance of the two-parameter Weibull distribution in describing the wind speed distribution, another two distributions: Logistic and Lognormal distributions are adopted. The pdfs of these two distributions are expressed in the following equations, respectively:

$$f_1(v) = \frac{\exp[-(v-\overline{v})/\delta]}{\delta\{1+\exp[-(v-\overline{v})/\delta]\}^2}, \tag{21}$$

$$f_2(v) = \frac{1}{v\phi\sqrt{2\pi}}\exp\left\{-\frac{[\ln(v)-\lambda]^2}{2\phi^2}\right\}, \tag{22}$$

where $\delta = \sqrt{3}\sigma/\pi$ is called the Logistic scale parameter, $\overline{v}$ and σ are estimated in this paper by the moment estimation approach, i.e., they are estimated by the mean wind speed and the standard deviation of the actual wind speed series. Similarly, λ and ϕ are estimated by the mean and the standard deviation of natural logarithm of the observed wind speed.

The cumulative distribution functions of the Logistic and the Lognormal distributions can be calculated by the integrations of Eqs. (21) and (22), which can be expressed as follows:

$$F_1(v) = \frac{1}{1+\exp[-(v-\overline{v})/\delta]} - \frac{1}{1+\exp(-\overline{v}/\delta)}, \tag{23}$$

$$F_2(v) = \frac{1}{2} + \frac{1}{2} * \frac{2}{\sqrt{\pi}}\int_0^x \exp(-t^2)dt, \tag{24}$$

where $x = \ln(v) - \overline{v}/\sqrt{2}\sigma$.

7.2. Evaluation criteria

The performance of the above three distributions is evaluated according to four criteria: the first is the mean relative error

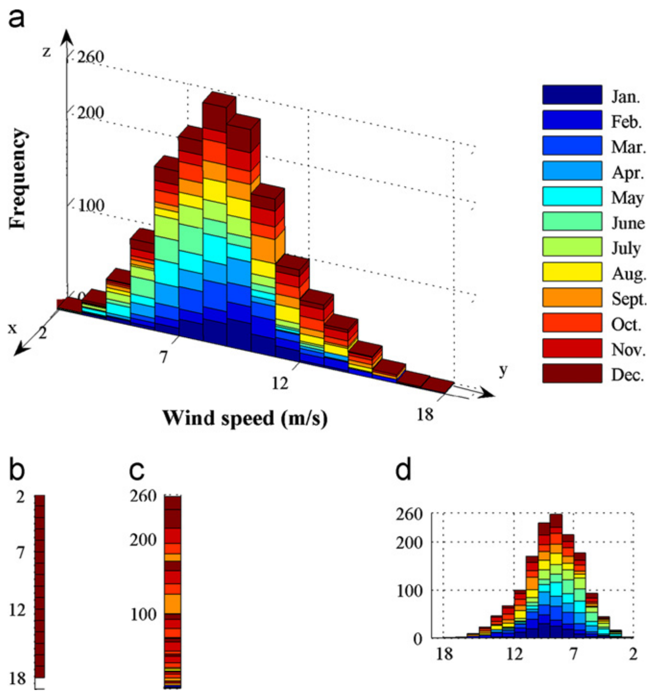


Fig. 4. Frequency of the yearly wind speed located in each wind speed interval: (a) three-dimensional view; (b) x–y view; (c) x–z view; (d) y–z view.

(MRE), which is given as follows:

$$\text{MRE} = \frac{1}{n}\sum_{i=1}^{n}\frac{|x_{oi}-x_{ci}|}{x_{oi}}, \tag{25}$$

where $\{x_{oi}\}_{i=1}^{n}$ and $\{x_{ci}\}_{i=1}^{n}$ are the observed and calculated data series respectively, and n is the number of the data. Besides this, the Chi-square error (CSE), defined in Eq. (26) is used as the second measure just as Liu et al. did in [34]:

$$\text{CSE} = \sum_{i=1}^{n}\frac{(x_{oi}-x_{ci})^2}{x_{ci}}. \tag{26}$$

As featured in Eq. (26), CSE is actually a relative error between $(x_{oi}-x_{ci})^2$ and $x_{ci}(i=1,2,\ldots,n)$, thus here the third criterion which can be used to measure the relative error between $(x_{oi}-x_{ci})^2$ and x_{oi} $(i=1,2,\ldots,n)$ is adopted. Because of its similarity with the CSE, it has been called the similar Chi-square error (SCSE):

$$\text{SCSE} = \sum_{i=1}^{n}\frac{(x_{oi}-x_{ci})^2}{x_{oi}}. \tag{27}$$

In addition, the root mean square error (RMSE) is also employed to evaluate the performance of different distributions:

$$\text{RMSE} = \left[\frac{1}{n}\sum_{i=1}^{n}(x_{oi}-x_{ci})^2\right]^{1/2}. \tag{28}$$

7.3. Simulation and comparison results

To analyze the wind speed distribution in the area described in Section 3 by the three distributions, i.e., Weibull, Logistic and Lognormal, the following procedures are employed in dividing wind speed ranges:

Step 1: Marking the minimum and the maximum wind speed data as MI and MA.

Step 2: Generating several intervals with the equal length 1 like this: $[Floor(MI), Floor(MI)+1)$, $[Floor(MI)+1, Floor(MI)+2)$, … $[Ceil(MA)-1, Ceil(MA))$, where $Ceil(MA)$ is the nearest integer larger than MA and $Floor(MI)$ is defined as itself if MI is an integer; otherwise, it is set as the nearest integer smaller than MI.

By the above range division method, the yearly wind speed data are divided into different intervals and the number of wind speed data located in each interval is calculated. Fig. 4 gives the frequency of the wind speed located in each wind speed interval for the twelve months throughout the year.

From Fig. 4(a) it can be observed that 16 intervals are generated in all: [2,3),[3,4), …[17,18). In addition, the wind speed data located in the interval [8,9) are the greatest, and in the interval [9,10) are the second greatest, while the wind speed data fall in range [17,18) are the least. Fig. 4(b) is the x–y view of Fig. 4(a). It shows that all of the 16 intervals contain the wind speed data in December. A y–z view of Fig. 4(a) is shown in Fig. 4(d). It can be observed that the maximum frequency of wind speed located in interval [8,9) is about 260, while the minimum in the range [17,18) is close to 0.

By the statistical number given in Fig. 4, the probability density function and the cumulative distribution function in each month can be easily obtained. Now applying the three different distributions mentioned in Section 6.1 to wind speed modeling, the pdf and cdf of these three distributions are obtained. Fig. 5 shows the monthly pdf and cdf of the observed wind speed and three distributions in four different months: February (in Winter), May (in Spring), August (in Summer) and November (in Autumn). As presented in Fig. 5(a)–(d), the probability for wind speed data located in ranges [10,11), [8,9), [8,9) and [9,10) are the largest for February, May, August and November, respectively. The maximum probability values calculated by the Logistic pdf for these months are all much closer to the observed probability values than those calculated by the other two distributions. For pdf lines in May displayed in Fig. 5(b), the observed and the calculated maximum probability values by the Logistic pdf are greater than 0.2, while for the other three months, the maximum probability values are smaller than 0.2. Additionally, the maximum wind speed in May is located in the interval [11,12), while the maximum wind speed data for the other three months are all larger, and the one in February which is located in the range [17,18) is the largest. In addition, the cdf lines plotted by the three distributions for February fit the actual cdf line better than for other months.

Apart from the monthly wind speed distribution modeling, the yearly wind speed distribution analysis is also quite meaningful. Thus Fig. 6 provides the pdf and cdf comparison between the observed wind speed data and data calculated from the three distributions. Similar to May and August, the probability for wind speed data located in the interval [8,9) is the maximum.

Since figures only provide a rough view to different models or distributions, thus the numerical calculations for the three distributions are provided to give a clear description. Error values between the observed and the calculated values for the 12 months and the whole year are listed in Table 5. As seen, there are several incalculable values when the MRE or SCSE is selected to be the performance evaluation criterion. This is because once zero value appears in the sequence $\{x_{oi}\}_{i=1}^{n}$, the definition of the MRE or SCSE makes no sense. Thus these two error criteria have drawbacks. However, for CSE, the same defect also exists when the zero value appears in the sequence $\{x_{ci}\}_{i=1}^{n}$, while the fourth evaluation criterion (RMSE) can be put to use in any cases.

Results in Table 5 reveal that the MRE obtained by the Weibull distribution is the smallest for May and July as compared to the other two distributions. Apart from February, April and December, in which the smallest MRE values are obtained by the Logistic distribution, for other months and the whole year, the Lognormal distribution performs best under this error evaluation criterion.

When CSE, SCSE or RMSE is selected to be the error evaluation criterion, the situation is similar: For some months, the Weibull distribution performs best in fitting the wind speed distribution, while for other months, the best distribution is the Logistic or the Lognormal distribution.

As a result, under different evaluation criteria and months, the three distributions perform discriminately. In Table 5, the three distributions are compared 13 times (12 months and the whole year) under each judgment criterion, thus for each time, three error values are obtained. Recording the minimum and the maximum value as '*S*' and '*L*', and the remaining one as '*M*', then the performance frequency of the three distributions under each evaluation criterion is calculated, the results are plotted in Fig. 7. As shown in the MRE of Fig. 7, the best performance frequency for Weibull, Logistic and Lognormal distributions are 2, 6 and 3 respectively. These values change to 5, 5 and 3 for CSE, 3, 5, 3 for SCSE and 3, 4, 6 for RMSE. Consequently, as compared to the Logistic and Lognormal distributions, the performance of the Weibull distribution is not as good as the one mentioned in [34], which used the Weibull function to analyze the wind speed distribution finally for its best ability. Thus, there is no consistent conclusion to tell which function is the best in modeling wind speed distribution. Function performance should be analyzed and understood under specific conditions. Optimal function should be determined according to the time, area resource of the sampled data or the evaluation criterion which has been taken into account preferentially. For example, if one prioritizes the RMSE, the Lognormal function will be the optimal; while if only SCSE and RMSE are considered and they occupy the same weight, the frequency of the appearance of 'S' appears for the three distributions should be computed firstly. The calculated values are 6 (3+3=6), 9 (5+4=9) and 9 (3+6=9) respectively, thus both of the Logistic and Lognormal distributions can be regarded as the best.

In this paper, considering the drawback of the MRE and the SCSE criterions in calculating the errors, these two criterions have

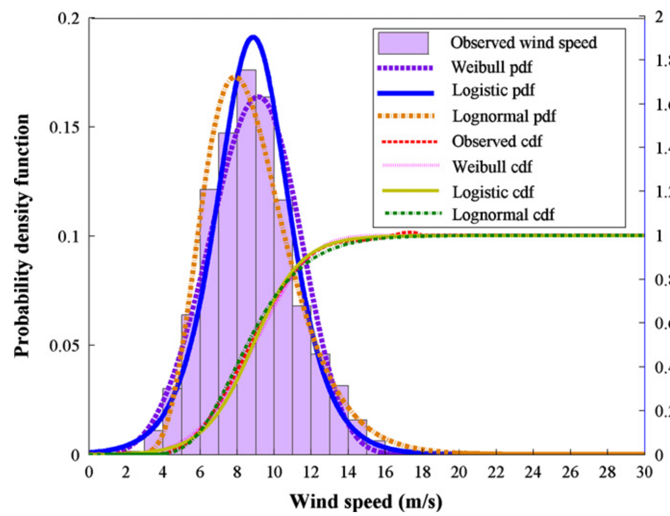


Fig. 6. Yearly pdf and cdf of the observed wind speed and three distributions.

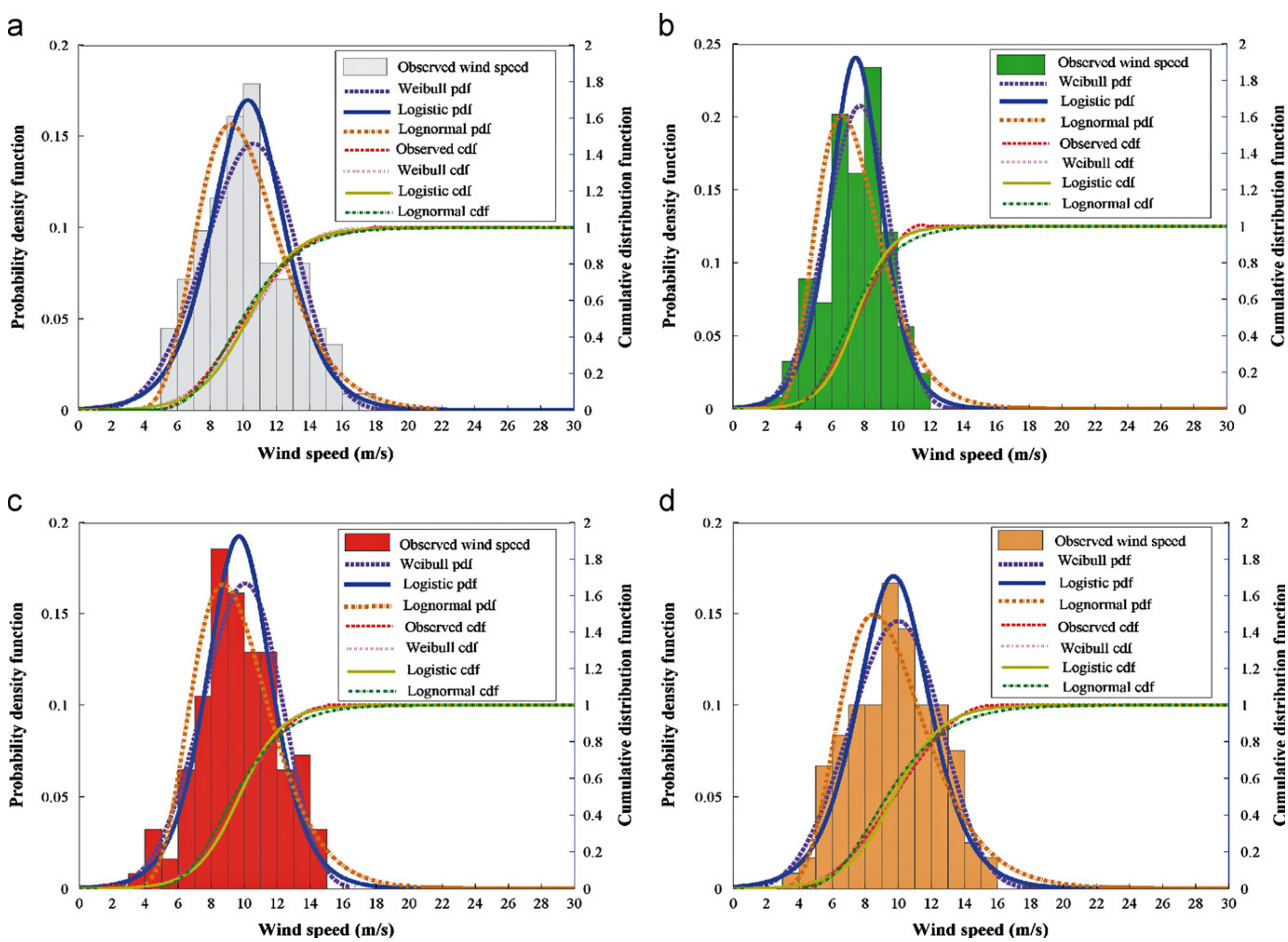


Fig. 5. Monthly pdf and cdf of the observed wind speed and three distributions in four different months: (a) February, (b) May, (c) August, (d) November.

been given lower weight coefficient than the other two criterions CSE and RMSE. The weight coefficients for MRE, CSE, SCSE and RMSE are 1/6, 1/3, 1/6 and 1/3 respectively. So the final performance of the three distributions is

Weibull: $1/6*2+1/3*5+1/6*3+1/3*3=7/2$;
Logistic: $1/6*6+1/3*5+1/6*5+1/3*4=29/6$;
Lognormal: $1/6*3+1/3*3+1/6*3+1/3*6=4$.

For $29/6>7/2$ and $29/6>4$, thus the Logistic distribution is the best distribution in fitting the yearly wind speed in this area, and the following analyses are all based on this best distribution.

8. Wind energy calculation and analysis

8.1. Wind energy calculation

The most probable wind speed for a particular location with the given probability function can be computed according to the calculation of the maximum value of the Logistic pdf by $\partial f(v)/\partial v=0$, by which the following formula can be obtained:

$$\frac{\exp[-(v-\overline{v})/\delta]\{\exp[-(v-\overline{v})/\delta]-1\}}{\delta^2\cdot\{1+\exp[-(v-\overline{v})/\delta]\}^3}=0, \tag{29}$$

thus the most probable wind speed can be expressed as

$$v_{MP}=\overline{v}. \tag{30}$$

The cumulative Logistic distribution function enables the calculation of the probability of the wind speed exceeding the value u:

$$F(v\geq u)=1-F(v<u)=1-\frac{1}{1+\exp[-(v-\overline{v})/\delta]}+\frac{1}{1+\exp(\overline{v}/\delta)}. \tag{31}$$

For a wind turbine designed with a given cut-in speed v_i and cut-off speed v_o, the operating probability of the wind turbine, named here availability factor, is equal to the probability of wind speed between v_i and v_o, calculated by

$$A_F=F(v_i\leq v<v_o)=F(v<v_o)-F(v<v_i)$$

Table 5
Comparison results of three distributions.

Period	MRE			CSE			SCSE			RMSE		
	Weibull	Logistic	Lognormal	Weibull	Logistic	Lognormal	Weibull	Logistic	Lognormal	Weibull	Logistic	Lognormal
January	0.3995	0.3353	0.6762	0.0627	0.0652	0.3609	0.0830	0.0565	0.2446	0.0190	0.0170	0.0395
February	0.2803	0.2881	0.2459	0.0952	0.0973	0.0667	0.0949	0.0977	0.0592	0.0242	0.0231	0.0200
March	0.3869	0.2197	0.2518	0.0670	0.0319	0.2075	0.0706	0.0309	0.0246	0.0229	0.0156	0.0113
April	1.0433	0.7923	0.7118	0.4016	0.2480	0.1922	0.5142	0.3037	0.2928	0.0384	0.0369	0.0363
May	0.2134	0.2512	0.3661	0.0535	0.0871	0.1774	0.0606	0.0888	0.1726	0.0272	0.0334	0.0434
June	– [b]	– [b]	– [b]	3.1385	0.1324	0.3173	– [b]	– [b]	– [b]	0.0302	0.0237	0.0144
July	0.1369	0.2040	0.2432	0.0252	0.0538	0.0692	0.0226	0.0487	0.0539	0.0178	0.0251	0.0228
August	0.3043	0.2850	0.3932	0.0767	0.0959	0.2118	0.0837	0.0702	0.0980	0.0218	0.0216	0.0211
September	0.4460	0.3611	0.3768	0.1738	0.1129	0.1719	0.1414	0.0858	0.1395	0.0270	0.0229	0.0313
October	– [b]	– [b]	– [b]	0.1246	0.1326	0.0659	– [b]	– [b]	– [b]	0.0271	0.0307	0.0180
November	0.2210	0.2079	0.2772	0.0372	0.0821	0.0969	0.0375	0.0639	0.0822	0.0149	0.0205	0.0244
December	0.5190	0.3746	0.3376	0.1070	0.0699	0.2020	0.1276	0.0745	0.0861	0.0234	0.0198	0.0275
Yearly	0.5227	0.4192	0.6245	0.0693	0.0294	0.0520	0.0552	0.0373	0.0454	0.0130	0.0103	0.0112

[b] Incalculable value.

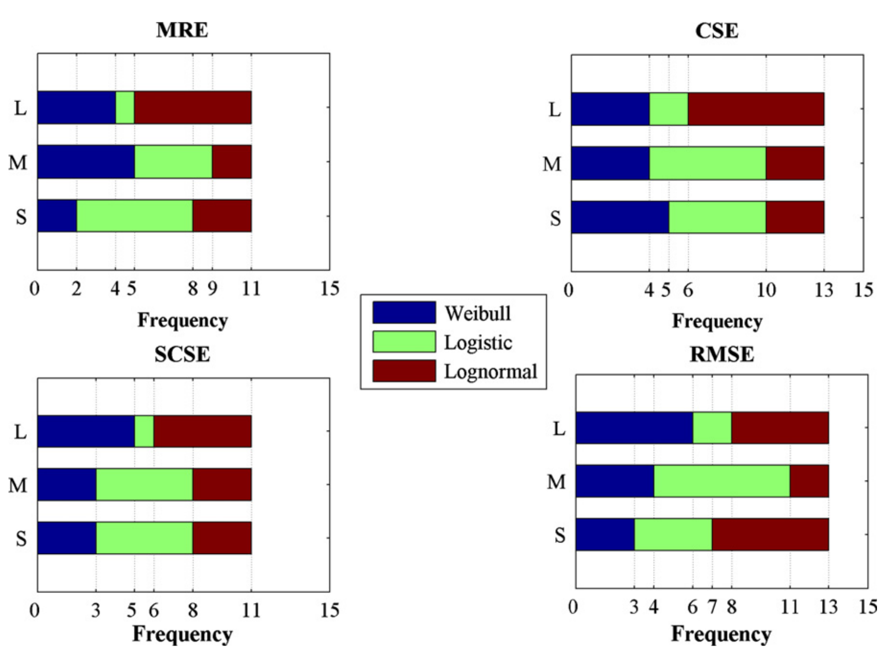


Fig. 7. Frequency of the three distributions under each judgment criterion.

$$= \frac{1}{1+\exp[(-\nu_o - \overline{\nu})/\delta]} - \frac{1}{1+\exp[-(\nu_i - \overline{\nu})/\delta]}. \quad (32)$$

8.2. Capacity factor

The capacity factor is a very significant index of productivity of a wind turbine. It is the ratio of the actual energy produced in a given period, to the hypothetical maximum possible, i.e. running full time at rated power. In this paper, the capacity factor is calculated by the following expression:

$$CF = \frac{E_W}{E_R}, \quad (33)$$

where

$$E_W = T\int_0^\infty P_{wtg}(\nu) f(\nu) d\nu, \quad (34)$$

$$P_{wtg}(\nu) = \begin{cases} P_r, & \nu_r < \nu \le \nu_o, \\ P_r \frac{\nu-\nu_i}{\nu_r-\nu_i}, & \nu_i < \nu \le \nu_r, \\ 0, & \nu \le \nu_i \text{ or } \nu > \nu_o, \end{cases} \quad (35)$$

$$E_R = TP_r. \quad (36)$$

ν_i, ν_r, ν_o, P_r and T represents the cut-in, rated, cut-off speed, the rated power and the time. In this paper, $f(\nu)$ is the probability density function of the Logistic function.

8.3. The turbine efficiency

The turbine efficiency η, which is defined as the ratio of the recoverable energy on the aerogenerator and the available energy at Betz limit, can be simply calculated by the following formula [37]:

$$\eta = \frac{t}{A} * \frac{P_r}{P_h}, \quad (37)$$

where t the time period, A is the blade sweep area of wind turbine, P_r is the rated power of the wind turbine and P_h is defined as

$$P_h = \frac{t}{1000}\frac{1}{2}\overline{\rho}\nu_m^3\left(\frac{h}{10}\right)^{3\alpha}, \quad (38)$$

in which ν_m is the mean wind speed of this time period at the standard height 10 m, h is the hub height, α is the roughness factor similar as Eq. (1), and the corrected air density $\overline{\rho}$ (kg/m^3) is determined by

$$\overline{\rho} = \frac{\overline{P}}{R_d * \overline{T}}, \quad (39)$$

where $\overline{P}$ and $\overline{T}$ are the mean air pressure and the mean air temperature in the time period t with the unit of N/m^2 and Kelvin, respectively and R_d=287.

8.4. Wind energy results and analysis

Wind and wind turbine characteristics for different time periods have been summarized in Table 6. As seen, all of the availability factors in this area are larger than 0.98. The capacity factor presents a similar trend with the mean wind speed for each time period: the capacity factor in February is the largest. Meanwhile, the mean wind speed in this month (see Table 1) is also the maximum. Similarly, the smallest capacity factor value and the minimum most probable speed appear in the same month (May). So in the month with the larger mean wind speed, the wind turbine could utilize more wind energy. Furthermore, the smallest value for the turbine efficiency is 0.6555. According to Betz, under the assumption that the turbine swirl and transmission losses were ignored, the theoretically maximum power that can be extracted from the wind is 59% of the wind power available in the wind. Therefore, for any wind turbine, the wind turbine efficiency should not exceed 0.59 [37]. Thus, the wind energy potential in this site of Inner Mongolia appears to be abundant.

To analyze the wind energy output, six other different types of wind turbine generators are adopted. The specifications of these turbines are given in Table 7, among which the maximum cut-in, rated and cut-off speeds are 5, 15 and 27 m/s, respectively. The maximum rated power is as high as 3000 kW. Meanwhile, two wind turbines (CONE-450 and YT/850) with the same hub height (50 m) with the given wind turbine are selected to survey the influence of different rated speed to the availability and capacity factors.

The availability and capacity factors for different wind turbines are listed in Table 8. As seen, the availability factors of the S43/600, CONE-450 and YT/850 are the same for all of the months and the whole year. This is because the cut-in and cut-off of these three wind turbines are the same as well as the hub height. It can be concluded that the availability factor is only relative to the cut-in and cut-off speed, but has no relation with the rated speed and the rated power. In addition, for wind turbines with the same cut-in and cut-off speed, the higher the hub height, the larger the availability factor. This can be viewed by the availability factor values of the NODEX-N 70/1500 and SL 3000/90 wind turbines. Capacity factor values for different wind turbines in Table 8

Table 7
Specifications of different wind turbines.

Number	Type	P_r/ kW	H/ m	ν_i/ (m/s)	ν_r/ (m/s)	ν_o/ (m/s)
1	CONE-450	450	50	3	12	25
2	NM600/43	600	56	4	15	25
3	YT/850	850	50	3	15	25
4	NODEX-N70/1500	1500	70	3	13	25
5	GAMMA-60	1500	66	5	13.3	27
6	SL 3000/90	3000	90	3	13	25

Table 6
Wind and wind turbine characteristics for different time periods.

Item	Period												
	January	February	March	April	May	June	July	August	September	October	November	December	Yearly
Logistic scale parameter (m/s)	1.0863	1.4730	1.0184	1.1528	1.0399	1.0504	1.1165	1.2996	1.3165	1.3791	1.4659	1.1522	1.3087
Most probable speed (m/s)	8.9439	10.2558	8.6957	8.2298	7.4566	7.5508	7.8993	9.6786	9.5821	9.5605	9.6972	8.7084	8.8444
Availability factor	0.9958	0.9927	0.9963	0.9894	0.9864	0.9870	0.9877	0.9942	0.9933	0.9915	0.9897	0.9930	0.9886
Capacity factor	0.4954	0.6006	0.4748	0.4366	0.3725	0.3803	0.4093	0.5554	0.5475	0.5455	0.5561	0.4760	0.4873
The corrected air density	1.1647	1.1686	1.1692	1.1772	1.1834	1.1898	1.1999	1.1975	1.1901	1.1824	1.1760	1.1695	1.1807
The turbine efficiency	0.9917	0.6555	1.0748	1.2594	1.6843	1.6133	1.3971	0.7611	0.7892	0.7998	0.7706	1.0699	1.0116

Table 8
Availability and capacity factors for different wind turbines.

Factor	Number	Period												
		January	February	March	April	May	June	July	August	September	October	November	December	Yearly
Availability factor	1	0.9958	0.9928	0.9963	0.9894	0.9864	0.9870	0.9877	0.9942	0.9933	0.9915	0.9897	0.9930	0.9886
	2	0.9901	0.9864	0.9907	0.9764	0.9672	0.9689	0.9720	0.9881	0.9864	0.9833	0.9807	0.9843	0.9770
	3	0.9958	0.9928	0.9963	0.9894	0.9864	0.9870	0.9877	0.9942	0.9933	0.9915	0.9897	0.9930	0.9886
	4	0.9963	0.9933	0.9968	0.9906	0.9881	0.9886	0.9891	0.9947	0.9939	0.9923	0.9906	0.9938	0.9897
	5	0.9782	0.9758	0.9785	0.9511	0.9273	0.9315	0.9409	0.9769	0.9740	0.9691	0.9656	0.9672	0.9562
	6	0.9966	0.9937	0.9971	0.9914	0.9892	0.9896	0.9901	0.9951	0.9944	0.9928	0.9911	0.9943	0.9905
Capacity factor	1	0.6539	0.7636	0.6290	0.5777	0.4953	0.5055	0.5428	0.7205	0.7107	0.7061	0.7150	0.6280	0.6385
	2	0.4630	0.5792	0.4402	0.3987	0.3282	0.3367	0.3687	0.5294	0.5207	0.5186	0.5304	0.4418	0.4548
	3	0.4954	0.6007	0.4748	0.4366	0.3725	0.3803	0.4093	0.5554	0.5475	0.5455	0.5561	0.4760	0.4873
	4	0.6333	0.7436	0.6091	0.5606	0.4820	0.4917	0.5273	0.6994	0.6899	0.6859	0.6953	0.6088	0.6197
	5	0.5170	0.6564	0.4873	0.4329	0.3407	0.3518	0.3938	0.5995	0.5884	0.5847	0.5977	0.4891	0.5049
	6	0.6645	0.7710	0.6401	0.5899	0.5094	0.5193	0.5558	0.7292	0.7196	0.7151	0.7236	0.6391	0.6492

illustrate that for each month and the whole year, the capacity factor of the SL 3000/90 is the largest, and the one of the CONE-450 is the second largest. The reason is that the capacity factor is influenced by the cut-in, rated and cut-off speeds, as well as the probability density function. This can be clearly seen by the simplification type of Eq. (33) shown as below:

$$CF = \int_{\nu_i}^{\nu_r} \frac{\nu - \nu_r}{\nu_r - \nu_i} f(\nu) d\nu + A_F. \tag{40}$$

9. Conclusions

In this paper, wind resource in one of the stations in Inner Mongolia, China, is analyzed. First, the wind speed probability distribution is modeled by three probability density functions: the Weibull, Logistic and Lognormal functions, among which the shape parameter in the Weibull function is estimated through three PSO algorithms and 18 DE approaches, and the one with the smallest minimum iteration number under the same fitness level is selected to determine the final Weibull shape parameter. Then four error evaluation criteria are adopted to choose the best probability density function by giving them different weight, and the best Logistic function is used in the consequent wind energy assessment. Capacity and availability factors are two necessary aspects in wind energy assessment. Thus, the two factors of six other different wind turbines are compared with the selected wind turbine. The conclusions are summarized as below:

(a) There is no function that performs best under any conditions in wind speed distribution modeling. So function performance should be analyzed and understood under specific conditions. Optimal function should be determined according to the time, area resource of the sampled data or the evaluation criterion judgment criterion which is taken into account preferentially.
(b) Performance of the intelligent parameter optimization methods, such as the PSO and DE, is not always better than that of the conventional moment estimation approach in searching the optimal parameter values.
(c) The available wind speeds in this area for each month and the whole year are larger than 98%.
(d) The capacity factor presents a similar trend with the mean wind speed for each time period, thus in the month with the larger mean wind speed, the wind turbine can utilize more wind energy.
(e) This area is suitable for a wind farm foundation.

Acknowledgment

The work was supported by the National Natural Science Foundation of China (Grant No.71171102). The authors are grateful to Mrs. Jan Pope for her assistance in editing the language of this paper.

References

[1] Department of Energy Statistics, National Bureau of statistics, People's Republic of China, China Energy Statistical yearbook, China statistics Press; 2011.
[2] Ling Y, Cai X. Exploitation and utilization of the wind power and its perspective in China. Renewable and Sustainable Energy Reviews 2012;16:2111–7.
[3] Energy Condition and Policy of China 〈http://baike.baidu.com/view/1332248.htm〉.
[4] Zhou Y, Wu WX, Liu GX. Assessment of onshore wind energy resource and wind-generated electricity potential in Jiangsu, China. Energy Procedia 2011;5:418–22.
[5] Zhou JY, Erdem E, Li G, Shi J. Comprehensive evaluation of wind speed distribution models: A case study for North Dakota sites. Energy Conversion and Management 2010;51:1449–58.
[6] Brano VL, Orioli A, Ciulla G, Culotta S. Quality of wind speed fitting distributions for the urban area of Palermo, Italy. Renewable Energy 2011;36:1026–39.
[7] Kiss P, Jánosi IM. Comprehensive empirical analysis of ERA-40 surface wind speed distribution over Europe. Energy Conversion and Management 2008;49:2142–51.
[8] Jaramillo OA, Borja MA. Wind speed analysis in La Ventosa, Mexico: a bimodal probability distribution case. Renewable Energy 2004;29:1613–30.
[9] Akpinar S, Akpinar EK. Wind energy analysis based on maximum entropy principle (MEP)-type distribution function. Energy Conversion and Management 2007;48:1140–9.
[10] Akpinar S, Akpinar EK. Estimation of wind energy potential using finite mixture distribution models. Energy Conversion and Management 2009;50:877–84.
[11] Kantar YM, Usta I. Analysis of wind speed distributions: Wind distribution function derived from minimum cross entropy principles as better alternative to Weibull function. Energy Conversion and Management 2008;49:962–73.
[12] Seguro JV, Lambert TW. Modern estimation of the parameters of the Weibull wind speed distribution for wind energy analysis. Journal of Wind Engineering and Industrial Aerodynamics 2000;85:75–84.
[13] Eskina N, Artara H, Tolun S. Wind energy potential of Gokceada Island in Turkey. Renewable and Sustainable Energy Reviews 2008;12:839–51.
[14] Pérez IA, García MA, Sánchez ML, Torre B. Analysis of height variations of sodar-derived wind speeds in Northern Spain. Journal of Wind Engineering and Industrial Aerodynamics 2004;92:875–94.
[15] Chang TP. Performance comparison of six numerical methods in estimating Weibull parameters for wind energy application. Applied Energy 2011;88:272–82.
[16] Weisser D. A wind energy analysis of Grenada: an estimation using the 'Weibull' density function. Renewable Energy 2003;28:1803–12.
[17] Akpinar EK, Akpinar S. Determination of the wind energy potential for Maden-Elazig, Turkey. Energy Conversion and Management 2004;45:2901–14.

[18] Keyhani A, Varnamkhasti MG, Khanali M, Abbaszadeh R. An assessment of wind energy potential as a power generation source in the capital of Iran, Tehran. Energy 2010;35:188–201.
[19] Ohunakin OS, Akinnawonu OO. Assessment of wind energy potential and the economics of wind power generation in Jos, Plateau State, Nigeria. Energy for Sustainable Development 2012;16:78–83.
[20] Chang TJ, Wu YT, Hsu HY, Chu CR, Liao CM. Assessment of wind characteristics and wind turbine characteristics in Taiwan. Renewable Energy 2003;28:851–71.
[21] Köse R. An evaluation of wind energy potential as a power generation source in Kütahya, Turkey. Energy Conversion and Management 2004;45:1631–41.
[22] Ullah I, Chaudhry QZ, Chipperfield AJ. An evaluation of wind energy potential at Kati Bandar, Pakistan. Renewable and Sustainable Energy Reviews 2010;14:856–61.
[23] Mirhosseini M, Sharifi F, Sedaghat A. Assessing the wind energy potential locations in province of Semnan in Iran. Renewable and Sustainable Energy Reviews 2011;15:449–59.
[24] Đurišić Ž, Mikulović J. Assessment of the wind energy resource in the South Banat region, Serbia. Renewable and Sustainable Energy Reviews 2012;16:3014–23.
[25] Mathew S, Pandey KP, Kumar AV. Analysis of wind regimes for energy estimation. Renewable Energy 2002;25:381–99.
[26] Amar FB, Elamouri M, Dhifaoui R. Energy assessment of the first wind farm section of Sidi Daoud, Tunisia. Renewable Energy 2008;33:2311–21.
[27] Onat N, Ersoz S. Analysis of wind climate and wind energy potential of regions in Turkey. Energy 2011;36:148–56.
[28] Durak M, Şen Z. Wind power potential in Turkey and Akhisar case study. Renewable Energy 2002;25:463–72.
[29] Hossain J, Sinha V, Kishore VVN. A GIS based assessment of potential for wind farms in India. Renewable Energy 2011;36:3257–67.
[30] Acker TL, Williams SK, Duque EPN, Brummels G, Buechler J. Wind resource assessment in the state of Arizona: Inventory, capacity factor, and cost. Renewable Energy 2007;32:1453–66.
[31] Thiaw L, Sow G, Fall SS, Kasse M, Sylla E, Thioye S. A neural network based approach for wind resource and wind generators production assessment. Applied Energy 2010;87:1744–8.
[32] Pepper DW, Wang XL. Application of an h-adaptive finite element model for wind energy assessment in Nevada. Renewable Energy 2007;32:1705–22.
[33] Oh KY, Kim JY, Lee JK, Ryu MS, Lee JS. An assessment of wind energy potential at the demonstration offshore wind farm in Korea. Energy 2012;46:555–63.
[34] Liu FJ, Chen PH, Kuo SS, Su DC, Chang TP, Yu YH, et al. Wind characterization analysis incorporating genetic algorithm: a case study in Taiwan Strait. Energy 2011;36:2611–9.
[35] Yan LP, Zeng JC. Particle swarm optimization with self-adaptive stochastic inertia weight. Computer Engineering and Design 2006;27:4677–9 in Chinese.
[36] Fang W, Sun J, Xu WB. Improved quantum-behaved particle swarm optimization algorithm based on differential evolution operator and its application. Journal of System Simulation 2008;20:6740–4 in Chinese.
[37] Ahmed AS. Wind energy as a potential generation source at Ras Benas, Egypt. Renewable and Sustainable energy review 2010;14:2167–73.